

# Metadata for You and Me: Moving Towards More Shareable Metadata

Sarah Shreeve  
University of Illinois at Urbana-Champaign  
WebWise 2006 February 16, 2006



?????



## Collection Registries



Cushman ID No.: 947.14  
Roll No.: 9-47  
Date: Aug. 10, 1947  
Location: Chicago, Illinois, United States (Cook county)  
Description: Circus Clown and black pig  
[View details](#)



Cushman ID No.: 947.15  
Roll No.: 9-47  
Date: Aug. 10, 1947  
Location: Chicago, Illinois, United States (Cook county)  
Description: Clown Felix Adler  
[View details](#)



Cushman ID No.: 947.16  
Roll No.: 9-47  
Date: Aug. 10, 1947  
Location: Chicago, Illinois, United States (Cook county)  
Description: Clown Felix Adler and bottle fed pig  
[View details](#)

Photograph from  
Indiana University  
Charles W. Cushman  
Collection



# Aggregations Aplenty

Metadata and content appearing in lots and lots of places – open web, subscription databases, specialized aggregations, general aggregations...

RLG Cultural Materials Online \* National Science Digital Library \* IMLS  
Digital Collections and Content \* New Jersey Digital Highway \* Sheet  
Music Consortium \* Digital Library Federation MODS Portal \* Online  
Archive of California \* CIC Metadata Portal \* Heritage Colorado \* Picture  
Australia \* Maine Memory Network \* American West \* OAlster \* Louisiana  
Digital Library \* Central Florida Memory \* Connecticut History Online \*  
Illinois Digital Archive \* Global Performing Arts Database ...

But what are we (aggregators) doing with that metadata?

Primarily simple aggregations

Primarily search and discovery

We need to build smart aggregations  
which requires that  
we make our data work harder  
(all credit to Lorcan Dempsey)

# Metadata from DLF MODS Portal in Amazon's A9 Search

## Web Results

[\[full\]](#) [\[close\]](#)

Showing 1 - 10 of about 39,400

### Cushman

Huge selection of **Cushman** items. Shop for fun recreational vehicles.  
[www.MonsterMarketplace.com](http://www.MonsterMarketplace.com)

### Books by Illinois

Illinois & more at Amazon  
Qualified orders over \$25 ship free  
[Amazon.com/books](http://Amazon.com/books)

Latest news for **illinois cushman** :)

- German firm seeks 70% of **Illinois** Center tower - Chicago Tribune - Feb 15, 2006

### CUSHMAN PAGE

Karen **Cushman** was born in Chicago, **Illinois**. As a child, she learned about the public library very early. She would go alphabetically down the shelves and ...

<http://www.scils.rutgers.edu/~kvander/cushman.html> - 15k

Cached [Site Info](#)

### Gapers Block, Chicago, IL -

Photo Essay: Chicago, Then &

## Image Results

[\[full\]](#) [\[close\]](#)

Showing 1 - 17 of 17



## Digital Library Federation MODS Portal

[\[full\]](#) [\[close\]](#)

1 - 10 of about 2,488

### Cairo, Illinois . Court House. Cairo, Ill.

**Title:** Cairo, **Illinois**. Court House. Cairo, Ill.  
**Alternative Title:** County [...] House.

**Author/Creator:** **Cushman**, Charles Weaver, 1896-1972, photographer, pht **Type:** [...] Fire hydrants **Subject:** **Hierarchical:** United States; **Illinois**; Alexander; Cairo Junction **Identifier:** 1952.14; P06185 [...] **URL:** <http://purl.dlib.indiana.edu/iudl/archives/cushman>

**Rights:** Copyright and reproduction rights for all [...] W. **Cushman** photographs are held by Indiana University and administered by the University [...] Indiana University Digital Library Program: Charles W. **Cushman** Photograph Collection [...]

### Cairo, Ill. Honey locust ? St. Mary's Park. Cairo, Ill.

[...] tree ? (with pods) in St. Mary's Pk.

**Author/Creator:** **Cushman**, Charles Weaver, 1896-1972, photographer, pht **Type:** [...] (Botany); Acacia **Subject:** **Hierarchical:** United States; **Illinois**; Alexander; Cairo Junction **Identifier:** 1952.13; P06184 [...] **URL:** <http://purl.dlib.indiana.edu/iudl/archives/cushman>

**Rights:** Copyright and reproduction rights for all [...] W. **Cushman** photographs are held by

# What can aggregators do to enable shareable metadata?

- Better ways to navigate aggregations
- Specialized aggregations
  - Content focused
  - Audience focused
  - Services focused
- Develop and share tools and services that make the data work harder
  - Metadata remediation and normalization
  - Clustering
  - Incorporating collection description and other contextual information

## Best Practices for Shareable Metadata

A MS Word version of this document is available at: <http://comm.nsdll.org/download.php/653/> (August 2005)

### Introduction

- ♦ [Introduction]

### General recommendations

- ♦ Metadata should be a [shareable and appropriate representation of the resource]
- ♦ [Granularity of description]
- ♦ Use of [metadata formats] in addition to simple Dublin Core is strongly encouraged
- ♦ [Tips for crosswalking]
- ♦ [Describing versions and reproductions]
- ♦ [Linking from a record to a resource] and other linking issues
- ♦ Providing [supplemental information] to service provider
- ♦ [Expressing rights over metadata]

### Recommendations for classes of data elements

- ♦ [Titles]
- ♦ [Names]
- ♦ [Dates]
- ♦ [Subjects / Topics]
- ♦ [Language]
- ♦ [Geographic Places]
- ♦ [Identifiers]
- ♦ [Rights for Resources]
- ♦ [Bibliographic Citations]
- ♦ [Types of Resources]

### Best Practices for Technical Aspects of Metadata

- ♦ [XML Namespaces and Schemas]
- ♦ [Character Encoding Issues]

### Final Preparations

- ♦ [Final preparations]

**Collaborative  
Effort  
to Distill the  
Experience  
of OAI  
service and  
data  
providers....**

*Supported by an IMLS  
grant to the Digital Library  
Federation*

# Shareable Metadata...

- Is quality metadata (see Bruce and Hillmann)
- Is (at the least) human understandable outside of its local context
- Preferably is machine processable!
- Is an appropriate view of the resource
- Promotes search interoperability
  - “the ability to perform a search over diverse sets of metadata records and obtain meaningful results.”  
(Priscilla Caplan)
- Is useful outside of its local context
  - (Can we build something off of it?)



# The Five C's (and lots of S's) of Shareable Metadata

Consistency

Coherence

Context

Communication

Conformance

Metadata standards

Vocabulary and encoding standards

Descriptive content standards

Technical standards

# What else?

- Expression of provenance and reliability (trust)
- Better understanding of what's worth putting effort into
- Metadata registries
- Giving back remediated metadata
- Demand that our vendors incorporate standards and give us the control to enable more shareable metadata (and content)



## A Last Word...

“We are beginning to explore how to share metadata for the digitized collections, and have very good technical solutions, but this has not yet matured into a well understood set of services. This is one aspect of what I meant when I said that this activity was in the 'cottage industry' stage.”

Lorcan Dempsey

March 5, 2005

<http://orweblog.oclc.org/archives/000602.html>



# Copyright notice

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 2.5 License.

To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/2.5/> or send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.